Nonlinear Estimation and Control of Particle Trajectories in the Ocean

PI: Toshio Michael Chin
CoPI: Ashwanth Srinivasan
Division of Meteorology and Physical Oceanography
Rosenstiel School of Marine and Atmospheric Science
4600 Rickenbacker Causeway, Miami, Florida 33149
phone: 305-421-4057 fax: 305-421-4696 email: tchin@rsmas.miami.edu

Award Number: N000140710175
http://www.rsmas.miami.edu/divs/mpo/People/Faculty/Chin/

LONG-TERM GOALS

Our long-range goal is to develop optimization methods: 1) to estimate the physical state of the ocean in order to understand the present and future conditions and associated variability/uncertainty, and 2) to utilize such forecast information for control-decisions such as optimal drifter deployment strategy. This is being accomplished through the use of data assimilation methods for ocean circulation models and the study of extending the assimilation formulation to an optimal control problem.

OBJECTIVES

In this effort, we study application of the Monte Carlo numerical techniques to problems of ocean data assimilation and optimal drifter deployment. This report focuses on our continuing efforts on application of the *particle filter* to the inverse Lagrangian prediction problem relevant to drifter deployment.

APPROACH

An *inverse Lagrangian prediction* (ILP) problem addresses retrospective estimation of drifter trajectories through a turbulent flow given their final positions. In a typical ILP scenario, the launch location of a single or a set of drifters in the past is sought given the present location(s) of these drifter(s). Due to chaotic nature of the forward Lagrangian problem and limitations in accuracy and resolution of current and wind data, it is usually difficult to expect an unique and deterministic answer to an ILP problem. It may, however, be possible to estimate the launch site and time statistically so that the drifters deployed in the estimated region and time would have the largest

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Nonlinear Estimation and Control of Particle Trajectories in the Ocean**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rosenstiel School of Marine and Atmospheric Science,Division of Meteorology and Physical Oceanography,4600 Rickenbacker Causeway,Miami,FL,33149**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

probability of arriving at the desired final location. For most practical problems it is desirable to minimize the optimal deployment region while maximizing the probability of successful delivery.

One approach to solving the ILP problem is to simulate an ensemble of Lagrangian trajectories backwards in time using the known final locations and a stochastic model of the flow field. Due to the typically fast rate (exponential or geometrical functions of time) dispersion of the trajectories, however, the distribution of the drifter locations tends to be too diffuse to be able to locate the launch site. We have investigated a numerical method that employs the particle filter to control the spread of the drifter location. The goal is to localize the optimal deployment region using the compact distribution of the constrained drifter positions.

WORK COMPLETED

1) The initial paper on application of particle filter to the ILP problem has been submitted and accepted for publication in *Journal of Atmospheric and Oceanic Technology* (Chin and Mariano, 2009). The article presents the methodology and the results from controlled experiments.

2) Extensions of the published work to (i) more realistic ILP scenarios and (ii) more strongly non-Gaussian cases have started; see RESULTS and IMPACT/APPLICATIONS sections below for more details.

3) Inter-comparison study involving the EnROIF assimilation system, a Monte-Carlo (or ensemble-based) enhancements to the existing ROIF method (Chin et al 2002), is completed for a 1/12-degree resolution HYCOM over Gulf of Mexico, and the result (Srinivasan et al, 2009) will be submitted for publication this calendar year.

RESULTS

We evaluate the benefit of the *resampled particle filter* (RPF; Chin et al 2007) by comparing two ensembles of trajectories. One is an ensemble produced with the RPF procedure denoted as $\mathbf{r}_n^{\mathrm{RPF}}$, $n = 1, \ldots, N$; the other is an ensemble of trajectories without any constraint and denoted as $\mathbf{r}_n^{\mathrm{Ens}}$. To compare the two ensembles, the launch site distribution estimated by each ensemble is used to initialize some *test drifters* for *forward* trajectory simulations. The target locations estimated by the test drifters can then be used to evaluate statistical accuracy in reproducing the known target locations $\mathbf{X}_m$, $m = 1, \ldots, M$.

Two skill scores are computed to compare the ensembles $\mathbf{r}_n^{\mathrm{RPF}}$ and $\mathbf{r}_n^{\mathrm{Ens}}$. By letting $G(\mathbf{X}_m)$ be the chance (in %) of the m^{th} target being reached by any of test drifters, we define the *coverage score* to be $\gamma \equiv \min_m G(\mathbf{X}_m)$. We also define the μ to be average chance (in %) of a test drifter to reach any of the targets. A higher μ value indicates that a drifter from the ensemble is less likely to miss a target and that the drifter destination is more likely to be focused near a target location. We hence call μ the *resolution score*.

For ILP, we assume knowledge of an empirical time characteristics of *ensemble spread*, quantified here by the standard deviation

$$D_{\mathbf{r}}(t) = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\|\mathbf{r}_n(t) - \bar{\mathbf{r}}(t)\|^2} \tag{1}$$

where $\mathbf{r}(t)$ is the unknown drifter trajectory, $\mathbf{r}_n(t)$ is the n^{th} sample of such trajectory by simulation, and $\bar{\mathbf{r}}(t) \equiv \sum_{n=1}^{N}\mathbf{r}_n(t)/N$ is the ensemble-mean of such samples. In ILP, the drifter trajectories simulated backward in time are expected to converge towards each other due to causality. For example, in our test cases, we have found $D_{\mathbf{r}}(t) \propto t^b$ for a constant $b \in [1.0, 2.0]$ by forward simulations. To formally express the information with which to constrain the backward trajectory ensemble, we let $\mathbf{s}(t)$ be a fictitious "noisy observation" of the unknown trajectory $\mathbf{r}(t)$

$$\mathbf{s}(t) \;=\; \mathbf{r}(t) + \mathbf{e}(t) \tag{2}$$

where $\mathbf{e}(t)$ is vector of random observation errors each with a known variance E^2. Assuming that $\mathbf{r}(t)$ and $\mathbf{e}(t)$ are uncorrelated, the variance of $\mathbf{s}(t)$ would become $D_{\mathbf{r}}(t)^2 + E^2$. Since the mean of $\mathbf{s}(t)$, or an observation of the mean trajectory, is not available, we estimate it in a bootstrapping fashion using the ensemble mean $\bar{\mathbf{r}}(t)$ of the on-going simulation. The probability density function (PDF) $p_{\mathbf{s}|\mathbf{r}}$ of the observation s conditioned on the unknown $\mathbf{r}$ is used by the particle filter algorithm to constrain the state trajectory. The specific form of $p_{\mathbf{s}|\mathbf{r}}$ we use is

$$p_{\mathbf{s}|\mathbf{r}}(\mathbf{x}|\bar{\mathbf{r}}, t) = \frac{1}{c}\exp\left[-\frac{1}{2}\left(\frac{\|\mathbf{x} - \bar{\mathbf{r}}\|^2}{D_{\mathbf{r}}^2 + E^2}\right)^F\right] \tag{3}$$

where c is a normalization constant and F is a constant parameter to control "flatness" of the PDF. For $F = 1$, $p_{\mathbf{s}|\mathbf{r}}$ would become a Gaussian PDF. We use $F = 3$ so that the PDF would have a relatively flat peak near its maximum. Choosing the larger value of F would give more equal importance (probability) to the ensemble members within a certain distance from the maximum, rather than favoring those in the immediate vicinity of the maximum. More complete details of the numerical procedure can be found in (Chin and Mariano, 2009).

The PDF given in (3) is still uni-modal. We are thus investigating the use of a multi-modal PDF for $p_{\mathbf{s}|\mathbf{r}}$ by clustering the ensemble locations and then computing the mean $\bar{\mathbf{r}}$ for each of the cluster. The resulting PDF is similar to a *Gaussian mixture*, or a normalized sum of multiple Gaussian PDFs, except that the parameter F may differ from 1. To automate clustering, we use the well-known *Expectation Maximazation (EM)* algorithm designed for the Gaussian mixture PDF. The only free-variable parameter in this algorithm is the number of clusters.

The new multi-modal PDF for $p_{\mathbf{s}|\mathbf{r}}$ is applied to an array deployment scenario using the surface velocity field obtained from a $1/12°$ HYCOM over the Gulf of Mexico. Figure 1 shows a sample velocity field and the target array configuration (red dots) with 49 gridded locations. In this scenario the target locations are to be reached in 10 days after deployment. The deployment region estimated using an unconstrained ensemble of drifter locations (Fig. 1, black contours) is larger than those estimated using constrained ensembles (Figs. 2-5; Table 1, second column). For the constrained ensembles, 1 to 4 clusters have been used (Figs. 2-5, respectively). The "*coverage*" performance scores are significantly different depending on the number of clusters used (Table 1, third column), where the 4-clustered ensemble is the only one with the complete coverage (with a score of 100) among the constrained cases. This demonstrates the importance of clustering (and more generally the use of multi-modal PDF) in application of the particle filter to the ILP problem. In this example, the RPF (4-cluster) solution resulted in a deployment region with approximately half the area compared to the unconstrained solution, while the efficiency of the deployment is increased by 22% (Table 1).

IMPACT/APPLICATIONS

We have explored a particle filter approach to solve the inverse Lagrangian prediction problem by an ensemble simulation of backward trajectories. The numerical experiments demonstrate that ensemble spread can be controlled using a constraint derived empirically and that the constrained solution leads to a spatially more compact estimate of the launch site. The constrained solution is thus more efficient than the unconstrained counterpart, while not compromising much effectiveness in delivery to the intended target sites. Due to high demands for shipping resources in drifter deployment, adopting the technique to actual operations and evaluating its benefits would be potential topics of future investigation.

To this end, we have initiated collaborations with research groups interested in fish larval dispersion (Cowen et al, 2008) and oil spill contingency planning (Bergueiro et al, 2008). The "drifters" in these cases are not exactly passive tracers. For example, fish larva can locally migrate towards more favorable physical and chemical environments, while sea surface oil can gradually evaporate into atmosphere. The dynamic models in these cases will hence be more complex than those examined in our work so far. Application of the particle filter to these cases would still be straightforward, due to the flexibility of the algorithm.

RELATED PROJECTS

The data assimilation (EnROIF) component of this project has been associated with the U.S. GODAE: Global Ocean Prediction with the Hybrid Coordinate Ocean Model (HYCOM), in collaboration with the HYCOM Consortium (http://hycom.rsmas.miami.edu).

The ILP solution methodology developed in this project is planned to be incorporated into the *Connectivity Modeling System* designed for the larval dispersal study (Cowen et al, 2008). In initial studies, HYCOM simulated velocity fields over the Intra-America Seas region are used.

REFERENCES

Bergueiro JR, Lindo D, Moreno S, Calvilla JM, Gomez JI, Gonzalez JA. Study of the geographical boundaries for the free use of dispersants. preprint (English translation), 2008.

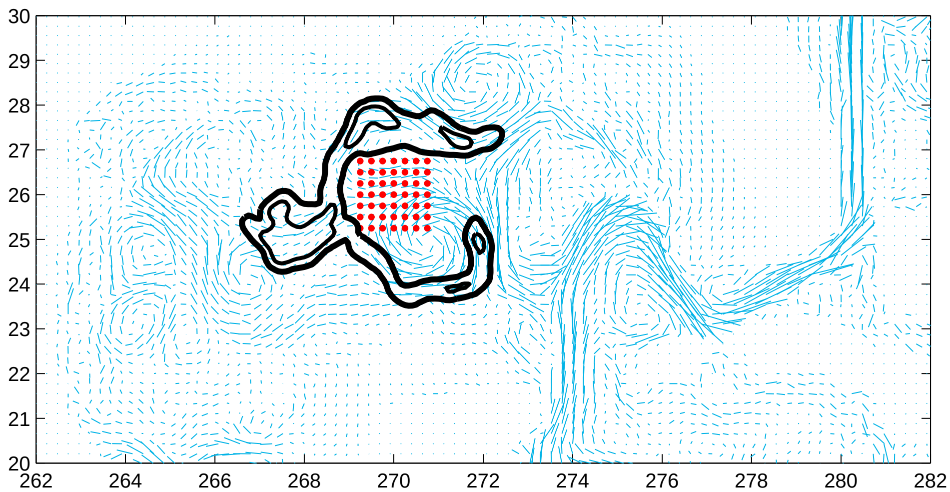

Figure 1. Unconstraint-ensemble. The black line is 95% probability contour of the prediction launch location to cover the target array (red dots) in the center of Gulf of Mexico. The prediction horizon is 10 days. The background flow (light blue vectors) is obtained from $1/12^{\circ}$ ***HYCOM output.***

Chin TM, Mariano AJ. A Particle Filter for Inverse Lagrangian Prediction Problem. Journal of Atmospheric and Oceanic Technology, in print, 2009.

Chin TM, Turmon MJ, Jewell JB, Ghil M. An ensemble-based smoother with retrospectively updated weights for highly nonlinear systems. *Mon. Wea. Rev.* 135: 186-202, 2007.

Chin TM, Haza AC, Mariano AJ. A reduced-order information filter for multi-layer shallow water models: profiling and assimilation of sea surface height. Journal of Atmospheric and Oceanic Technology 19: 517-533, 2002.

Cowen RK, Paris CB, Srinivasan A. Scaling of connectivity in marine populations. *Science* 311: 522-527, 2006.

Srinivasan A, Chassignet EP, Brasseur P, Chin TM, Counillon F, Cummings J, Smedstad OM, Thacker WC, Mariano AJ. A demonstration and comparison of sequential assimilation methods for HYCOM. Ocean Modelling, to be submitted, 2009.

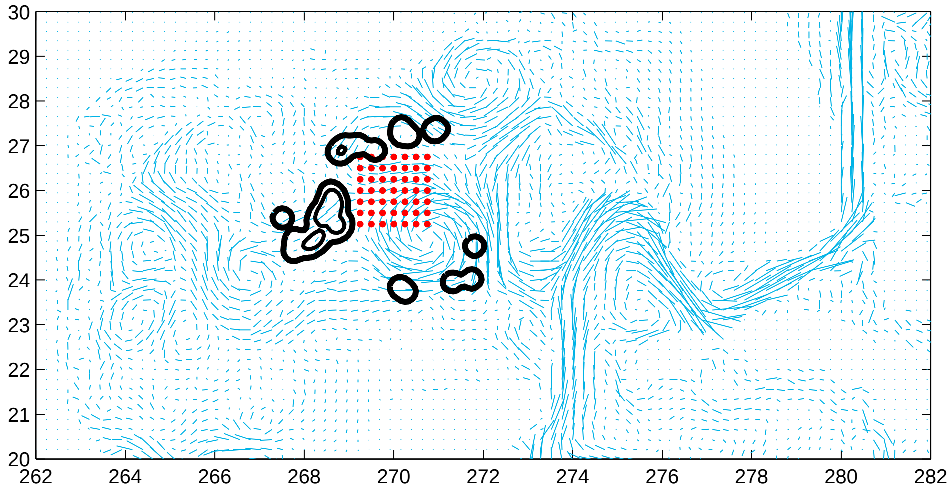

Figure 2. ***1-cluster RPF. The same as Fig. 1, except that the ensemble is constrained by an RPF employing $p_{\mathbf{s}|\mathbf{r}}$ with 1 cluster.***

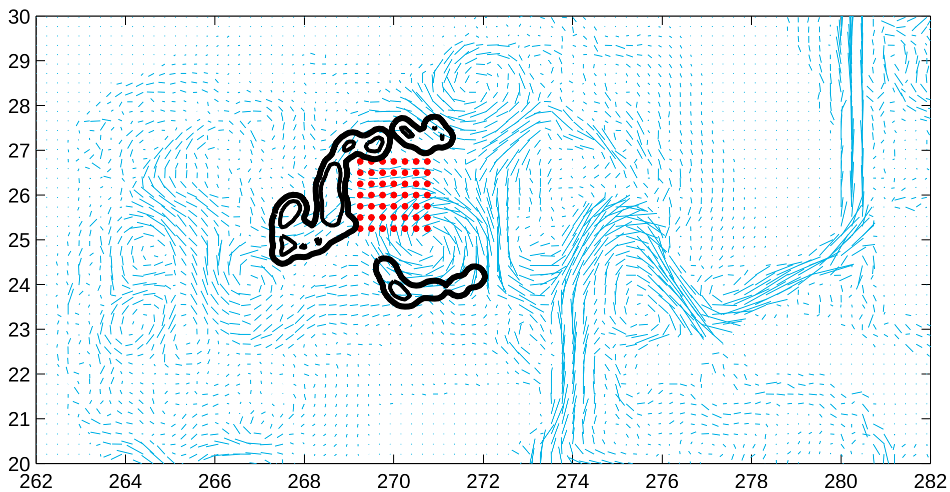

Figure 3. ***2-cluster RPF. The same as Fig. 1, except that the ensemble is constrained by an RPF employing $p_{\mathbf{s}|\mathbf{r}}$ with 2 clusters.***

Table 1. ***Skill scores from the array depolyment experiment.***

source site	release area $\left(A_{95}^{\mathrm{RPF}}/A_{95}^{\mathrm{Ens}}\right)$	coverage $(\min\gamma)$	efficiency $\left(\mu^{\mathrm{RPF}}/\mu^{\mathrm{Ens}}\right)$
Unconstrained Ens	1.00	100.0	1.00
1-cluster RPF	0.41	0.0	1.12
2-cluster RPF	0.60	1.6	1.05
3-cluster RPF	0.46	17.9	1.20
4-cluster RPF	0.51	100.0	1.22

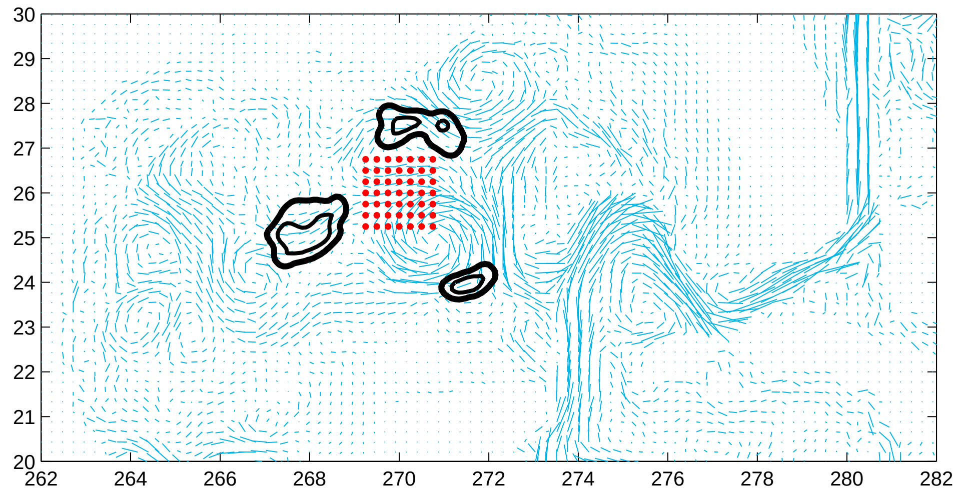

Figure 4. 3-cluster RPF. The same as Fig. 1, except that the ensemble is constrained by an RPF employing $p_{\mathbf{s}|\mathbf{r}}$ with 3 clusters.

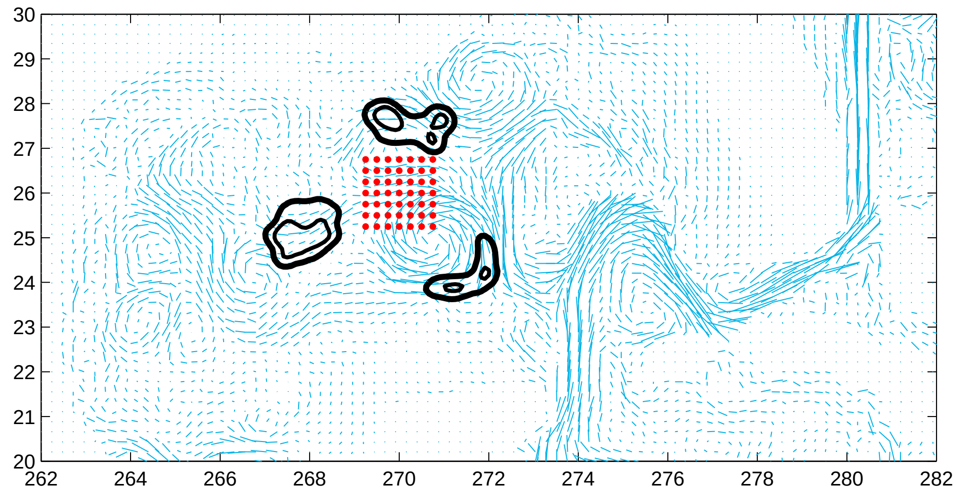

Figure 5. 4-cluster RPF. The same as Fig. 1, except that the ensemble is constrained by an RPF employing $p_{\mathbf{s}|\mathbf{r}}$ with 4 clusters.